Conical Averagedness and Fixed Point Algorithms

Hung Phan
Mathematical Sciences
University of Massachusetts Lowell

November 16, 2023

Based on joint work with Sedi Bartz and Minh N. Dao

Definitions

Let X be a Hilbert space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper lsc convex function.
The *subdifferential* of f at x: $\partial f(x) = \Big\{\text{all subgradients of } f \text{ at } x\Big\}$, where a vector u is called a *subgradient* of f at x if

$$\forall y \in X, \quad f(y) \geq f(x) + \langle u, y - x \rangle .$$

The *indicator function* of a set $\Omega \subset X$ is $\iota_\Omega(x) := \begin{cases} 0, & \text{if } x \in \Omega, \\ +\infty, & \text{otherwise.} \end{cases}$

The subdifferential of ι_Ω is the normal cone operator of Ω

$$\partial(\iota_\Omega)(x) = N_\Omega(x) = \Big\{u \in X, \langle u, z - x \rangle \leq 0, \forall z \in \Omega\Big\}$$

Fermat's Stationary

Let $f, g : X \to \mathbb{R} \cup \{+\infty\}$ be proper lsc convex functions. The Fermat's stationary condition:

$$\bar{x} \ \text{ solves } \ \min_{x \in X} f(x) \quad \Longleftrightarrow \quad 0 = \nabla f(\bar{x}) \quad (f \text{ is differentiable})$$

$$\bar{x} \ \text{ solves } \ \min_{x \in X} f(x) \quad \Longleftrightarrow \quad 0 \in \partial f(\bar{x}) \quad (\text{f is not differentiable})$$

$$\bar{x} \ \text{ solves } \ \min f(x) + g(x) \quad \Longleftarrow \quad 0 \in \partial f(\bar{x}) + \partial g(\bar{x})$$

$$\bar{x} \ \text{ solves } \ \min_{x \in \Omega} f(x) \quad \Longleftrightarrow \quad \bar{x} \ \text{ solves } \ \min f(x) + \iota_\Omega(x) \quad \Longleftarrow \quad 0 \in \partial f(\bar{x}) + N_\Omega(\bar{x})$$

$$\bar{x} \in \Omega_1 \cap \Omega_2 \quad \Longleftrightarrow \quad \bar{x} \ \text{ solves } \ \min_{x \in X} \iota_{\Omega_1}(x) + \iota_{\Omega_2}(x) \quad \Longleftarrow \quad 0 \in N_{\Omega_1}(\bar{x}) + N_{\Omega_2}(\bar{x})$$

So we may consider the inclusion problem: find an x such that

$$0 \in Ax + Bx \quad \text{where } A, B : X \rightrightarrows X \text{ are set-valued operators.}$$

Sum of Finitely Many Operators

Consider the problem

$$\min_{x \in X} f(x) \quad \text{s.t.} \quad x \in \Omega_2 \cap \cdots \cap \Omega_m,$$

which is equivalent to

$$\min_{x \in X} f(x) + \iota_{\Omega_2}(x) + \iota_{\Omega_3}(x) + \cdots + \iota_{\Omega_m}(x)$$

So we may consider solving the inclusion problem

$$0 \in A_1 x + A_2 x + \cdots + A_m x.$$

(where A_i's are the subdifferential operators of the functions involved).

Let $\mathbf{x} := (x_1, \ldots, x_m) \in X^m$. Define

$$\boldsymbol{A}(x) := A_1 x_1 \times \cdots \times A_m x_m$$

$$\text{and} \quad \boldsymbol{B}(\mathbf{x}) := N_\Delta(\mathbf{x}) \quad \text{where} \quad \Delta := \Big\{(x, \ldots, x) \in X^m\Big\}.$$

Then

$$0 \in A_1 x + A_2 x + \cdots + A_m x \quad \Longleftrightarrow \quad 0 \in \boldsymbol{A}(\mathbf{x}) + \boldsymbol{B}(\mathbf{x}).$$

Resolvent and Relaxed Resolvent

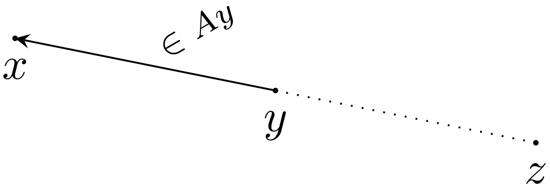
Let $A : X \rightrightarrows X$ be an operator.

The resolvent of A is defined by $\quad J_A := (\mathrm{Id} + A)^{-1}$

The reflected resolvent of A is defined by $\quad R_A := J_A^2 = 2J_A - \mathrm{Id}$

Let $\lambda > 0$, the λ-relaxed resolvent of A is defined by $\quad J_A^\lambda := (1-\lambda)\,\mathrm{Id} + \lambda J_A$

$$y \in J_A x \iff y = (\mathrm{Id} + A)^{-1} x \iff x \in (\mathrm{Id} + A) y = y + Ay$$



The resolvent of the (convex) normal cone operator $N_\Omega : X \rightrightarrows X$ is the projection:

$$J_{N_\Omega}(x) = P_\Omega(x) = \left\{ y \in \Omega\,,\ \|x - y\| = \min_{z \in \Omega} \|x - z\| \right\}$$

The Forward–Backward Algorithm

Consider the inclusion problem

$$0 \in Ax + Bx \quad \text{where } A \text{ is single-valued.}$$

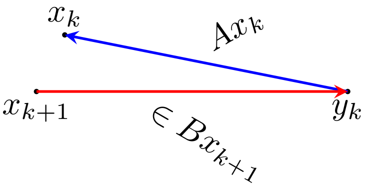
Consider the inclusion problem

$$0 \in Ax + Bx \quad \text{where } A \text{ is single-valued.}$$

Forward–Backward Algorithm:

Given x_k, define

$$y_k = x_k - Ax_k,$$
$$x_{k+1} = J_B y_k$$



- If $x_{k+1} = x_k = \bar{x}$, then $-A\bar{x} \in B\bar{x}$, so

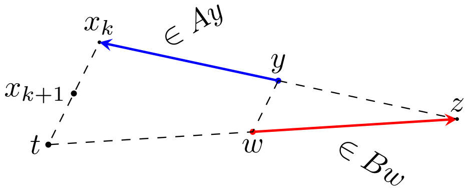
$$0 \in A\bar{x} + B\bar{x}.$$

The Douglas–Rachford (DR) Algorithm

$$x_{k+1} = Tx_k \quad \text{where} \quad T = \frac{1}{2}(\mathrm{Id} + R_B R_A).$$

Illustration:

$$y = J_A x_k, \quad z = R_A x, \quad w = J_B z, \quad t = R_B z, \quad x_{k+1} = \frac{1}{2}(x_k + t).$$



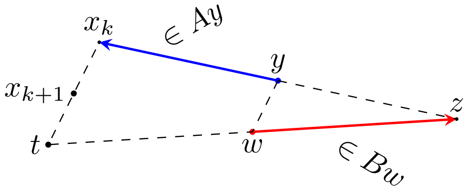
- If $x_{k+1} = x_k$, then $y = w$ and $0 \in Ay + Bw$. i.e., y is a solution.

The Douglas–Rachford (DR) Algorithm

$$x_{k+1} = Tx_k \quad \text{where} \quad T = \frac{1}{2}(\mathrm{Id} + R_B R_A).$$

Illustration:

$$y = J_A x_k, \quad z = R_A x, \quad w = J_B z, \quad t = R_B z, \quad x_{k+1} = \frac{1}{2}(x_k + t).$$



- If $x_{k+1} = x_k$, then $y = w$ and $0 \in Ay + Bw$. i.e., y is a solution.

Next, we will discuss conditions of A and B in order for these methods to converge.

Monotonicity, Firm Nonexpansiveness, and Nonexpansiveness

An operator A is monotone if $\forall (a,u),(b,v) \in \operatorname{gr} A, \quad \langle a-b, u-v\rangle \geq 0$.

A is maximally monotone if there is no monotone operator $\hat{A}$ such that $\operatorname{gr} A \subsetneq \operatorname{gr} \hat{A}$.

An operator T is nonexpansive (on its domain) if for all $x, y \in \operatorname{dom} T$, $\|Tx - Ty\| \leq \|x-y\|$.

An operator T is firmly nonexpansive (on its domain) if for all $x, y \in \operatorname{dom} T$,

$$\|Tx - Ty\|^2 \leq \|x-y\|^2 - \|(\mathrm{Id}-T)x - (\mathrm{Id}-T)y\|^2$$

Monotonicity, Firm Nonexpansiveness, and Nonexpansiveness

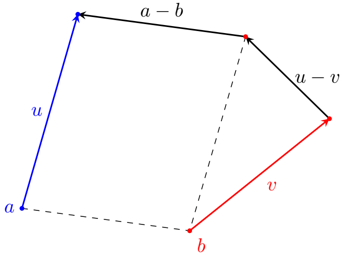
An operator A is monotone if $\forall (a,u),(b,v) \in \operatorname{gr} A, \quad \langle a-b, u-v\rangle \geq 0$.

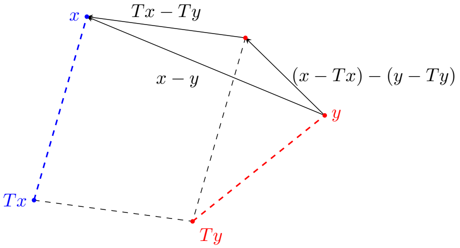
A is maximally monotone if there is no monotone operator $\hat{A}$ such that $\operatorname{gr} A \subsetneq \operatorname{gr} \hat{A}$.

An operator T is nonexpansive (on its domain) if for all $x, y \in \operatorname{dom} T$, $\|Tx - Ty\| \leq \|x - y\|$.

An operator T is firmly nonexpansive (on its domain) if for all $x, y \in \operatorname{dom} T$,

$$\|Tx - Ty\|^2 \leq \|x-y\|^2 - \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$$





A is monotone $\iff$ $T = (\mathrm{Id} + A)^{-1}$ is firmly nonexpansive

Monotonicity, Firm Nonexpansiveness, and Nonexpansiveness

Theorem (Minty's Theorem)

Let A be monotone. Then A is maximally monotone if and only if $\operatorname{dom}(\operatorname{Id}+A)^{-1} = X$.

Theorem

T is firmly nonexpansive if and only if $S := 2T - \operatorname{Id}$ is nonexpansive.

Theorem (Krasnosel'skiĭ–Mann)

Let $D \subset X$ and let $S : D \to D$ be a nonexpansive operator such that $\operatorname{Fix} S \neq \varnothing$. Let $\lambda_k \in [0,1]$ be such that $\sum_{k=1}^{\infty} \lambda_k(1-\lambda_k) = +\infty$ and $x_0 \in D$. Set

$$x_{k+1} = (1-\lambda_k)x_k + \lambda_k S x_k.$$

Then $(x_k)_{k\in\mathbb{N}}$ converges weakly to a fixed point of S.

- Let $0 < \lambda < 1$. We say that an operator T is λ-averaged if $T = (1-\lambda)\operatorname{Id} + \lambda S$ for some nonexpansive operator S. So if T is a λ-averaged operator and $\operatorname{Fix} T \neq \varnothing$, then the sequence $x_{k+1} = Tx_k$ converges weakly to a fixed point of T.

Convergence Analysis

Theorem ([Lions-Mercier 1979])

Let $A, B : X \rightrightarrows X$ be two maximally monotone operators such that $\operatorname{zer}(A + B) \neq \varnothing$. *Let* (x_k) *be a sequence generated by the Douglas–Rachford algorithm. Then* x_k *converges weakly to a fixed point* $\bar{x} \in \operatorname{Fix} T = \operatorname{Fix} R_B R_A$ *and* $J_A \bar{x} \in \operatorname{zer}(A + B)$.

Proof.

A and B are monotone

$\implies$ J_A and J_B are firmly nonexpansive

$\implies$ $R_A = 2J_A - \mathrm{Id}$ and $R_B = 2J_B - \mathrm{Id}$ are nonexpansive

$\implies$ $R_B R_A$ is nonexpansive

$\implies$ Apply the Krasnosel'skiĭ–Mann Theorem: x_k converges weakly to a fixed point. □

Theorem ([Svaiter '11] and [Bauschke '13])

The sequence $J_A x_k$ *converges weakly to* $J_A \bar{x}$.

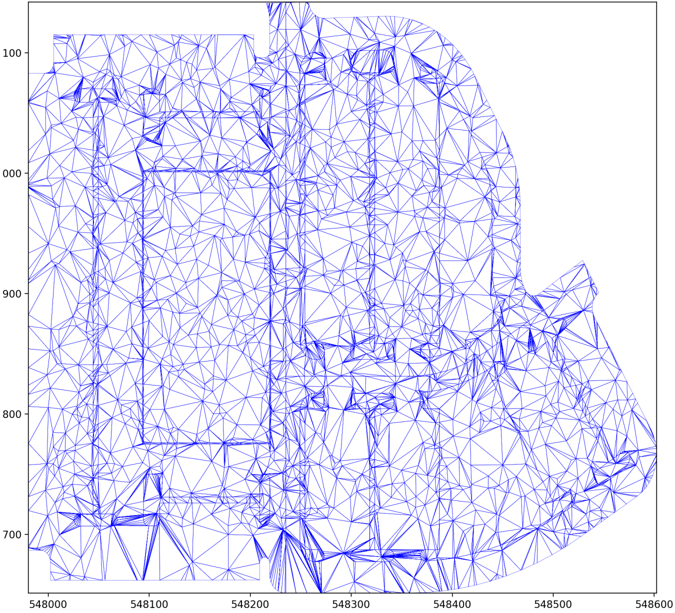
An Application (Base on the collaboration with Valentin Koch, Autodesk, Inc.)

A common problem in civil engineering design is the grading of a parking lot or a building pad. Within a given area, the engineer has to define grading slopes such that

- the grading site fits with existing structures.
- the drainage requirements on the surface are met.
- safety and comfort are taken into account.
- the engineer would like to change the existing surface as little as possible, in order to save on earthwork costs.

The grading site is usually represented as a Triangulated Irregular Network (TIN). The engineer is able to *adjust the heights* of the vertices in the triangulated grid, so that the newly obtained mesh-grid satisfies the above requirements.

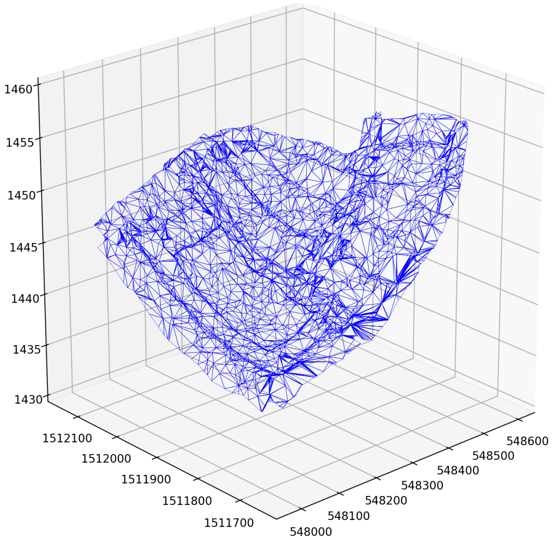
2D View of a Construction Site



$\#$ of vertices $\approx 5,000$

$\#$ of triangles: $\approx 7,000$

3D View of a Construction Site



The Triangular Mesh:

$$V = \{p_j = (p_{j1}, p_{j2}, z_j) \in \mathbb{R}^3\}, \ |V| = n,$$
$$E \subset \left\{ p_i p_j \,\middle|\, p_i, p_j \in V \right\},$$
$$T \subset \left\{ p_i p_j p_k \,\middle|\, p_i p_j, p_j p_k, p_k p_i \in E \right\}.$$

The variables are the **elevations** of the vertices, written as a vector

$$z = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$$

Conical Averagedness

Let $\theta > 0$, we say that an operator $T : X \to X$ is conically θ-averaged if

$$T = (1-\theta)\operatorname{Id} + \theta N \quad \text{for some nonexpansive operator } N.$$

$$\begin{array}{lcl} \theta = 1 & : & T \text{ is nonexpansive} \\ \theta = \frac{1}{2} & : & T \text{ is firmly nonexpansive} \\ \theta \in (0,1) & : & T \text{ is } \theta\text{-averaged} \end{array}$$

- Convex combinations of conically averaged operators are also conically averaged. In particular, if T is conically θ-averaged, then $(1-\kappa)\operatorname{Id} + \kappa T$ is conically $\kappa\theta$-averaged.

Theorem (Krasnosel'skiĭ–Mann)

Let $D \subset X$ and let $T : D \to D$ be a conically θ-averaged operator such that $\operatorname{Fix} T \neq \varnothing$. Let $\lambda \in (0, \frac{1}{\theta})$ and $x_0 \in D$. Set

$$x_{k+1} = ((1-\lambda)\operatorname{Id} + \lambda T)x_k = (1-\lambda)x_k + \lambda T x_k.$$

Then $(x_k)_{k\in\mathbb{N}}$ converges to a fixed point of T.

Theorem ([Bartz-Dao-Ph. '22])

Let $T_1, T_2 : X \to X$ be conically θ_1-averaged and conically θ_2-averaged. Suppose that either $\theta_1 = \theta_2 = 1$ or $\theta_1\theta_2 < 1$. Let also $\omega \in \mathbb{R} \smallsetminus \{0\}$. Then

$$T := \left(\frac{1}{\omega} T_2\right)(\omega T_1) \quad \textit{is conically } \theta\textit{-averaged with} \quad \theta := \begin{cases} 1, & \theta_1 = \theta_2 = 1, \\ \frac{\theta_1+\theta_2-2\theta_1\theta_2}{1-\theta_1\theta_2}, & \theta_1\theta_2 < 1. \end{cases}$$

In addition, if either $\theta_1 > 1$ or $\theta_2 > 1$, then $\theta > 1$.

Compositions of Conically Averaged Operators

Theorem ([Bartz-Dao-Ph. '22])

Let T_i be conically θ_i-averaged for $i = 1, \dots, m$ $(m \geq 2)$. Let $\omega_i \in \mathbb{R}$ be such that $\prod_{i=1}^m \omega_i = 1$. Set $T = (\omega_m T_m) \cdots (\omega_1 T_1)$. Then

(i) *If $\max_i \theta_i \leq 1$, then T is nonexpansive.*

(ii) *If $\theta_i \neq 1$ for all i and*

$$\theta_k < 1 + \frac{1}{\sum_{i=1}^{k-1} \frac{\theta_i}{1-\theta_i}},$$

Then T is conically θ-averaged with

$$\theta := \frac{1}{1 + \frac{1}{\sum_{i=1}^{m} \frac{\theta_i}{1-\theta_i}}}$$

(iii) *If $\max_i \theta_i < 1$, then T is θ-averaged with $\theta < 1$ given above.*

Generalized Monotonicity

Let $A : X \rightrightarrows X$ and $\alpha \in \mathbb{R}$. We say that A is

$$\begin{aligned} \alpha\text{-monotone if} \quad & \forall (x,u),(y,v) \in \operatorname{gr} A, \quad \langle x-y, u-v \rangle \geq \alpha \|x-y\|^2, \\ \alpha\text{-comonotone if} \quad & \forall (x,u),(y,v) \in \operatorname{gr} A, \quad \langle x-y, u-v \rangle \geq \alpha \|u-v\|^2, \end{aligned}$$

and maximally α-monotone/comonotone if there is no α-monotone/comonotone operator whose graph strictly contains $\operatorname{gr} A$.

- $\alpha = 0$: monotone.
- $\alpha > 0$: strongly monotone / strongly comonotone (= cocoercive).
- $\alpha < 0$: weakly monotone/ weakly comonotone.
- A is α-monotone iff A^{-1} is α-comonotone.

Generalized Monotonicity and Conical Averagedness [Bartz-Dao-Ph. '22]

Theorem (α-comonotone operators)

Let $\alpha, \lambda > 0$. The following are equivalent

(i) *A is α-comonotone (i.e., α-cocoercive).*

(ii) $\mathrm{Id} - \lambda A$ *is conically* $\frac{\lambda}{2\alpha}$*-averaged.*

Theorem (α-comonotone operators)

Let A be α-comonotone and suppose $\gamma + \alpha > 0$. Let $\lambda > 0$. Then

(i) *$J_{\gamma A}$ is conically* $\frac{\gamma}{2(\gamma+\alpha)}$*-averaged.*

(ii) *$R = (1-\lambda)\,\mathrm{Id} + \lambda J_{\gamma A}$ is conically* $\frac{\lambda\gamma}{2(\gamma+\alpha)}$*-averaged.*

Theorem (α-monotone operators)

Let A be α-monotone and let $\gamma > 0$ be such that $1 + \gamma\alpha > 0$. Then

(i) *$J_{\gamma A}$ is $(1+\gamma\alpha)$-comonotone.*

(ii) $\frac{1}{1-\lambda} R$ *is conically* $\frac{\lambda}{2(\lambda-1)(1+\gamma\alpha)}$*-averaged.*

Forward-Backward Algorithms Revisited [Bartz-Dao-Ph. '22]

The (relaxed) forward-backward algorithm: given $\gamma > 0, \kappa > 0$,

$$x_{n+1} = Tx_n \quad \text{where} \quad T = (1-\kappa)\,\mathrm{Id} + \kappa J_{\gamma A}(\mathrm{Id} - \gamma B)$$

Theorem

Suppose A is maximally α-comonotone, B is β-comonotone with $\beta > 0$. Suppose either

(i) $\alpha + \beta = 0$ *and* $\gamma = 2\beta$; *or*

(ii) $\alpha + \beta > 0$ *and* $\max\{0, 2\beta - 2\sqrt{\beta(\alpha+\beta)}\} < \gamma < 2\beta + 2\sqrt{\beta(\alpha+\beta)}$.

Then $T = (1-\kappa)\,\mathrm{Id} + \kappa J_{\gamma A}(\mathrm{Id} - \gamma B)$ *is conically averaged.*

Consequently, if $\mathrm{zer}(A+B) \neq \varnothing$ *and* κ *is appropriately chosen, then every sequence generated by* T *converges weakly to some fixed point in* $\mathrm{zer}(A+B)$.

Forward-Backward Algorithms Revisited [Bartz-Dao-Ph. '22]

The (relaxed) forward-backward algorithm: given $\gamma > 0, \kappa > 0$,

$$x_{n+1} = Tx_n \quad \text{where} \quad T = (1-\kappa)\operatorname{Id} + \kappa J_{\gamma A}(\operatorname{Id} - \gamma B)$$

Theorem

Suppose A is maximally α-comonotone, B is β-comonotone with $\beta > 0$. Suppose either

(i) $\alpha + \beta = 0$ *and* $\gamma = 2\beta$; *or*

(ii) $\alpha + \beta > 0$ *and* $\max\{0, 2\beta - 2\sqrt{\beta(\alpha+\beta)}\} < \gamma < 2\beta + 2\sqrt{\beta(\alpha+\beta)}$.

Then $T = (1-\kappa)\operatorname{Id} + \kappa J_{\gamma A}(\operatorname{Id} - \gamma B)$ *is conically averaged.*

Consequently, if $\operatorname{zer}(A+B) \neq \varnothing$ *and* κ *is appropriately chosen, then every sequence generated by* T *converges weakly to some fixed point in* $\operatorname{zer}(A+B)$.

Theorem

If A is maximally monotone, B is β-comonotone, $\beta > 0$, and $\gamma \in (0, 4\beta)$, then T is conically averaged.

Forward-Backward Algorithms Revisited [Bartz-Dao-Ph. '22]

The (relaxed) forward-backward algorithm: given $\gamma > 0, \kappa > 0$,

$$x_{n+1} = Tx_n \quad \text{where} \quad T = (1-\kappa)\,\mathrm{Id} + \kappa J_{\gamma A}(\mathrm{Id} - \gamma B)$$

Theorem

Suppose A is maximally α-comonotone, B is β-comonotone with $\beta > 0$. Suppose either

(i) $\alpha + \beta = 0$ *and* $\gamma = 2\beta$; *or*

(ii) $\alpha + \beta > 0$ *and* $\max\{0, 2\beta - 2\sqrt{\beta(\alpha+\beta)} < \gamma < 2\beta + 2\sqrt{\beta(\alpha+\beta)}$.

Then $T = (1-\kappa)\,\mathrm{Id} + \kappa J_{\gamma A}(\mathrm{Id} - \gamma B)$ *is conically averaged.*

Consequently, if $\mathrm{zer}(A+B) \neq \varnothing$ *and* κ *is appropriately chosen, then every sequence generated by* T *converges weakly to some fixed point in* $\mathrm{zer}(A+B)$.

Theorem

If A is maximally monotone, B is β-comonotone, $\beta > 0$, and $\gamma \in (0, 4\beta)$, then T is conically averaged.

- The classical convergence analysis for the forward-backward algorithm requires $\gamma \in (0, 2\beta)$.

The Adaptive Douglas–Rachford Algorithm (aDR)

Problem: Solve

$$0 \in Ax + Bx$$

where A and B are maximally α- and β- monotone with $\alpha + \beta \geq 0$; or
A and B are maximally α- and β- comonotone with $\alpha + \beta \geq 0$.

The Adaptive Douglas–Rachford Algorithm (aDR)

Problem: Solve

$$0 \in Ax + Bx$$

where A and B are maximally α- and β- monotone with $\alpha + \beta \geq 0$; or
A and B are maximally α- and β- comonotone with $\alpha + \beta \geq 0$.

- If A is α-monotone and B is β-monotone with $\alpha + \beta \geq 0$, then

$$A + B = \left(A - \frac{\alpha - \beta}{2}\,\mathrm{Id}\right) + \left(B + \frac{\alpha - \beta}{2}\,\mathrm{Id}\right) =: \widetilde{A} + \widetilde{B}.$$

Here, $\widetilde{A}$ and $\widetilde{B}$ are both $\left(\frac{\alpha+\beta}{2}\right)$-monotone, in particular, monotone.

So, one can simply solve the problem $0 \in \widetilde{A}x + \widetilde{B}x$ using classical tools.

The Adaptive Douglas–Rachford Algorithm (aDR)

Problem: Solve

$$0 \in Ax + Bx$$

where A and B are maximally α- and β- monotone with $\alpha + \beta \geq 0$; or
A and B are maximally α- and β- comonotone with $\alpha + \beta \geq 0$.

- If A is α-monotone and B is β-monotone with $\alpha + \beta \geq 0$, then

$$A + B = \left(A - \frac{\alpha - \beta}{2}\,\mathrm{Id}\right) + \left(B + \frac{\alpha - \beta}{2}\,\mathrm{Id}\right) =: \widetilde{A} + \widetilde{B}.$$

Here, $\widetilde{A}$ and $\widetilde{B}$ are both $\left(\frac{\alpha+\beta}{2}\right)$-monotone, in particular, monotone.

So, one can simply solve the problem $0 \in \widetilde{A}x + \widetilde{B}x$ using classical tools.

- We, however, examine the possibility of an algorithm on A and B!

The aDR: Formulation

$$x_{k+1} = Tx_k \quad , \quad T_{A,B} = (1-\kappa)\,\mathrm{Id} + \kappa R_2 R_1,$$

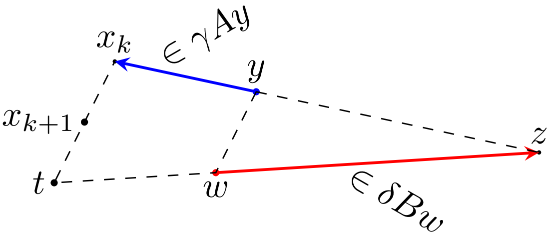
where $J_1 := J_{\gamma A}$, $J_2 := J_{\delta B}$

$$R_1 := (1-\lambda)\,\mathrm{Id} + \lambda J_1,\ R_2 := (1-\mu)\,\mathrm{Id} + \mu J_2,$$

$$\gamma > 0,\ \delta > 0,\ (\lambda-1)(\mu-1) = 1,\ \delta = \gamma(\lambda-1),\ \kappa \in\,]0,1[.$$

Illustration:

$$y = J_1 x_k\ ,\ z = R_1 x_k,\ w = J_2 z\ ,\ t = R_2 z,$$
$$x_{k+1} = (1-\kappa)x_k + \kappa t.$$



If $x_{k+1} = x_k \in \mathrm{Fix}\, T$, then

$$y = w \text{ and } 0 \in Ay + Bw,$$

i.e., y is a solution.

- If $\lambda = \mu = 2$, $\gamma = \delta > 0$, then the adaptive DR becomes the classical DR.

The aDR: Two Monotone Operators

Theorem ([Bartz-Dao-Ph. '22])

Assume A, B are maximally α-monotone and maximally β-monotone, $1 + 2\gamma\alpha > 0$, $\mu > 1$, and

$$\alpha + \beta \geq 0 \quad \text{and} \quad 2 + 2\gamma\alpha - \varepsilon \leq \mu \leq 2 + 2\gamma\alpha + \varepsilon \quad \text{with} \quad \varepsilon = 2\sqrt{\gamma(1 + \gamma\alpha)(\alpha + \beta)},$$

and either three strict inequalities happen simultaneously or none of them happens. Define

$$\lambda = \frac{\mu}{\mu - 1} \quad , \quad \delta = \frac{\gamma}{\mu - 1} \quad , \quad 0 < \kappa < \kappa^*,$$

where

$$\kappa^* := \begin{cases} 1, & \alpha + \beta = 0, \\ \frac{4\gamma\delta(1+\gamma\alpha)(1+\delta\beta) - (\gamma+\delta)^2}{2\gamma\delta(\gamma+\delta)(\alpha+\beta)}, & \alpha + \beta > 0. \end{cases}$$

Then the aDR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^*}$-averaged.

Consequently, if $(x_k)_{k\in\mathbb{N}}$ is a sequence generated by the aDR algorithm, then (x_k) converges weakly to a fixed point.

The aDR: Two Comonotone Operators

Theorem ([Bartz-Dao-Ph. '22])

Assume A, B are maximally α-comonotone and maximally β-comonotone, $\gamma + 2\alpha > 0$, and

$$\alpha + \beta \geq 0 \quad \text{and} \quad \gamma + 2\alpha - \varepsilon \leq \delta \leq \gamma + 2\alpha + \varepsilon \quad \text{with} \quad \varepsilon = 2\sqrt{(\gamma + \alpha)(\alpha + \beta)},$$

and either three strict inequalities happen simultaneously or none of them happens. Define

$$\lambda = 1 + \frac{\delta}{\gamma} \quad , \quad \mu = 1 + \frac{\gamma}{\delta} \quad , \quad 0 < \kappa < \kappa^*,$$

where

$$\kappa^* := \begin{cases} 1, & \alpha + \beta = 0, \\ \frac{4(\gamma+\alpha)(\delta+\beta) - (\gamma+\delta)^2}{2(\gamma+\delta)(\alpha+\beta)}, & \alpha + \beta > 0. \end{cases}$$

Then the aDR operators $T_{A,B}$ and $T_{B,A}$ are conically $\frac{\kappa}{\kappa^*}$-averaged.

Consequently, if $(x_k)_{k \in \mathbb{N}}$ is a sequence generated by the aDR algorithm, then (x_k) converges weakly to a fixed point.

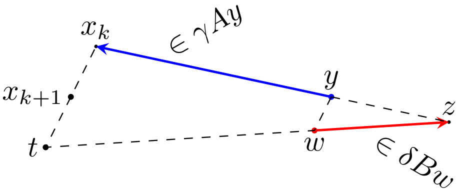
Remark: Under- and Over-Reflecting the Resolvents

Let $\alpha > 0$ and suppose that A is maximally α-monotone ("strong"),
B is maximally $(-\alpha)$-monotone ("weak").

Then

$$\mu = 2 + 2\gamma\alpha > 2 \quad \text{and} \quad \lambda = \frac{\mu}{\mu - 1} < 2.$$

- Under-reflect the resolvent of the strongly monotone operator A (use $\lambda < 2$).
- Over-reflect the resolvent of the weakly monotone operator B (use $\mu > 2$).



The Sum of Three Operators

Problem: Solve

$$0 \in Ax + Bx + Cx$$

where $A, B : X \rightrightarrows X$ are (generalized) monotone,
$C : X \to X$ is positively comonotone (i.e., cocoercive),
Resolvents of A and B are available,
Resolvents of C might not be available.

The Sum of Three Operators

Problem: Solve
$$0 \in Ax + Bx + Cx$$
where $A, B : X \rightrightarrows X$ are (generalized) monotone,
$C : X \to X$ is positively comonotone (i.e., cocoercive),
Resolvents of A and B are available,
Resolvents of C might not be available.

- A, B are maximally monotone: [Davis-Yin '17] proposes the fixed-point operator
$$T := \mathrm{Id} - J_A + J_B(2J_A - \mathrm{Id} - CJ_A).$$

One has
$$J_A(\mathrm{Fix}\, T) = \mathrm{zer}(A + B + C).$$

The Sum of Three Operators

Consider the fixed-point operator

$$T := \mathrm{Id} - \eta J_{\gamma A} + \eta J_{\delta B}((1-\lambda)\,\mathrm{Id} + \lambda J_{\gamma A} - \delta C J_{\gamma A})$$

where $\eta > 0$, $\gamma > 0$, $\delta > 0$, and $\lambda = 1 + \frac{\delta}{\gamma}$. Then

$$J_{\gamma A}(\mathrm{Fix}\, T) = \mathrm{zer}(A + B + C).$$

Theorem ([Dao-Ph. '21])

Suppose A, B are maximally α- and β-monotone with $\alpha + \beta = 0$ and C is σ-cocoercive with $\sigma > 0$. Suppose $\gamma > 0$ and $\eta > 0$ satisfy

$$1 + 2\gamma\alpha > 0 \quad \text{and} \quad \eta^* := 2 + 2\gamma\alpha - \frac{\gamma}{2\sigma} > 0$$

Set $\delta = \frac{\gamma}{1+2\gamma\alpha}$. Then the operator T is conically $\frac{\eta}{\eta^*}$-averaged.

- [Dao-Ph. '21] also includes a result for the case $\alpha + \beta > 0$.

Thank you!

- Bartz, S., Dao, M.N., Phan, H.M.: Conical averagedness and convergence analysis of fixed point algorithms, *Journal of Global Optimization* 82, 351–373 (2022)

- Bartz, S., Campoy, R., Phan, H.M.: Demiclosedness principles for generalized nonexpansive mappings, *J. Optim. Theory Appl.* **186**(3), 759–778 (2020)

- Bauschke, H.H.: New demiclosedness principles for (firmly) nonexpansive operators. In: Bailey D. et al. (eds) *Computational and Analytical Mathematics*, Springer Proceedings in Mathematics & Statistics. Springer, New York (2013)

- Dao, M.N., Phan, H.M.: An adaptive splitting algorithm for the sum of two generalized monotone operators and one cocoercive operator, *Fixed Point Theory and Algorithms for Sciences and Engineering* (2021)

- Dao, M.N., Phan, H.M.: Adaptive Douglas–Rachford splitting algorithm for the sum of two operators, *SIAM J. Optim.* **29**(4), 2697–2724 (2019)

- V.R. Koch and H.M. Phan, Optimization of triangular networks with spatial constraints, *Optim. Methods Softw.* (2019), DOI:10.1080/10556788.2019.1604703.

Thank you!

- Dao, M.N., Phan, H.M.: An adaptive splitting algorithm for the sum of two generalized monotone operators and one cocoercive operator, *Fixed Point Theory and Algorithms for Sciences and Engineering* (2021)
- Davis, D., Yin, W.: A three-operator splitting scheme and its optimization applications *Set-Valued Var. Anal.* 25(4), 829–858 (2017)
- Douglas, J., Rachford, H.H.: On the numerical solution of the heat condition problem in 2 and 3 space variables, *Trans. AMS.* **82**, 421–439 (1956)
- Guo, K., Han, D., Yuan, X.: Convergence analysis of Douglas–Rachford splitting method for "strongly + weakly" convex programming, *SIAM J. Numer. Anal.* **55**(4), 1549–1577 (2017).
- Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.* **16**, 964–979 (1979)
- Svaiter, B.F.: On weak convergence of the Douglas–Rachford method. *SIAM J. Control Optim.* **49**(1), 280–287 (2011)

Notes